

Durham Research Online

Deposited in DRO:

07 March 2019

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Aznan, N.K.N. and Connolly, J.D. and Al Moubayed, N. and Breckon, T.P. (2019) 'Using variable natural environment brain-computer interface stimuli for real-time humanoid robot navigation.', in 2019 International Conference on Robotics and Automation (ICRA) ; proceedings. , pp. 4889-4895.

Further information on publisher's website:

<https://doi.org/10.1109/ICRA.2019.8794060>

Publisher's copyright statement:

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Using Variable Natural Environment Brain-Computer Interface Stimuli for Real-time Humanoid Robot Navigation

Nik Khadijah Nik Aznan^{1,3}, Jason D. Connolly², Noura Al Moubayed¹ and Toby P. Breckon^{1,3}

Abstract—This paper addresses the challenge of humanoid robot teleoperation in a natural indoor environment via a Brain-Computer Interface (BCI). We leverage deep Convolutional Neural Network (CNN) based image and signal understanding to facilitate both real-time object detection and dry-Electroencephalography (EEG) based human cortical brain bio-signals decoding. We employ recent advances in dry-EEG technology to stream and collect the cortical waveforms from subjects while they fixate on variable Steady State Visual Evoked Potential (SSVEP) stimuli generated directly from the environment the robot is navigating. To these ends, we propose the use of novel variable BCI stimuli by utilising the real-time video streamed via the on-board robot camera as visual input for SSVEP, where the CNN detected natural scene objects are altered and flickered with differing frequencies (10Hz, 12Hz and 15Hz). These stimuli are not akin to traditional stimuli - as both the dimensions of the flicker regions and their on-screen position changes depending on the scene objects detected. On-screen object selection via such a dry-EEG enabled SSVEP methodology, facilitates the on-line decoding of human cortical brain signals, via a specialised secondary CNN, directly into teleoperation robot commands (approach object, move in a specific direction: right, left or back). This SSVEP decoding model is trained via *a priori* offline experimental data in which very similar visual input is present for all subjects. The resulting classification demonstrates high performance with mean accuracy of 85% for the real-time robot navigation experiment across multiple test subjects.

I. INTRODUCTION

Teleoperation or telepresence is a field within robotics which has been widely utilised for numerous applications. It allows humans to remotely control robots, either whilst being present within the same location, or remotely via the internet [1]. In this work, a humanoid robot is used as teleoperational remote control interface, allowing a human to navigate the robot via the use of Brain-Computer Interface (BCI) based cortical brain bio-signals [2]. This application can be used widely, for example by severely disabled people as an alternative communication platform with the robot without any actual physical movement [3].

A Brain-Computer Interface is a system that provides a communication and control medium between human cortical signals and external devices [4]. One of the primary aims of BCI is to assist or to be used by patients with Complete Locked-In Syndrome in which the end user cannot move or communicate due to paralysis, yet is cognitively intact and can therefore make real, tangible and informed decisions [2].

In order to gather the cortical signals from a human test subject, a non-invasive dry-Electroencephalography (EEG) will be used. EEG is a technique where electrodes or sensors are placed on the scalp to capture electrical activity of the brain, without the need to implant them directly into the brain, such as invasive microelectrode arrays [5]. We utilise the Cognionics Quick-20 dry-EEG Headset that requires no conductive gel and has the additional benefit of being a wireless device; as compared to traditional wet-EEG [6], [7]. Wet-EEG requires the cumbersome application of conductive gel, use of an expensive Faraday cage enclosure that prohibits real-world application and scratching of the skin surface via semi-invasive blunted needles in order to lower the impedance values - these measurements represent how usable the connectivity is between the electrodes and the scalp [8]. It is an alternative approach used to improve the usability of EEG within a BCI context via the elimination of these factors that is required for the wet-EEG approach [9].

This work explores the creation of a BCI-based application to accurately navigate a humanoid robot in an open environment via the above noted dry-EEG headset. We make use of NAO, a humanoid robot which is equipped with cameras and programmable movement and behavioural features such that different commands can be interpreted to navigate the robot to move toward the object of participant visual and cognitive interest [10]. To develop this effectively, we employ the features available from the humanoid robot such as streaming the real-time video from the robot as visual input for the SSVEP stimuli. SSVEP is a type of stimulus-evoked neurophysiological response induced simply via subject fixation (or even just via peripheral attention) on visual stimuli and requires almost no *a priori* user training [11]–[13]. The human cortical signals in the primary visual areas oscillate when visually evoked via these stimuli by a continuously fluctuating sinusoidal cycle [14], [15].

In this paper, we propose a novel variable dry-EEG enabled BCI stimuli for robot navigation utilising a pre-trained object detection convolutional neural network. We perform object detection in real-time derived from the incoming video stream from the robot camera. Our key idea is to make the SSVEP stimuli more natural to the user, as the stimuli (or in our case, objects) will be presented in the context of the real-world scene the robot is currently navigating. Unlike previous stimuli [6], [11], [16], in this work the size of each SSVEP flicker region depends on the physical dimensions of the object detected. The detected object pixel regions are flickered at differing on screen frequencies (10, 12, 15 Hz) and the decoded EEG signals are used to navigate the robot

¹Department of Computer Science, Durham University, Durham, UK.

²Department of Psychology, Durham University, Durham, UK.

³Department of Engineering, Durham University, Durham, UK.

to walk toward objects based on the objects selected by the subject (user) via the SSVEP interface.

To perform the dry-EEG signal decoding we use CNN architecture, detailed here [17], to differentiate between EEG signals by extracting unique features across multiple layers of convolutional transformation optimised over a set of training data [18]. This model is used to classify real-time dry-EEG signals before sending the decoded command to navigate the robot towards the scene object the subject has selected [19].

Following standard practice in the BCI literature, we evaluate the performance of our work by testing the classifier model on real-time humanoid navigation via classification accuracy and Information Transfer Rate (ITR) as performance metrics - the latter representing a quantitative measure of the speed of BCI information transfer [2].

In summary, the major contributions of this paper are:

- Use of a novel variable position and size SSVEP BCI stimuli, based on using object detection pixel regions identified in real-time, within the live video stream from a teleoperated humanoid robot traversing a real-world natural environment.
- An offline dry-EEG enabled SSVEP BCI signal decoding (classification) result achieving mean accuracy of 84% with the use of variable stimuli size and on-screen stimuli positioning (the first such study to accomplish this).
- Demonstrable real-time BCI teleoperation of a humanoid robot, based on the use of naturally occurring in-scene stimuli, with a peak mean accuracy of 85% and ITR of 15.2 bits per minute (bpm) when evaluated over multiple test subjects (teleoperation users).

II. RELATED WORK

There have been many prior studies utilising humanoid robots with EEG signals for various BCI applications [1], [4], [15]. In this section we will focus on the studies making use of SSVEP within this context.

The work of [10] proposed the use of behaviour-based SSVEP to control a telepresence humanoid robot to walk in a cluttered environment, with the tasking of approaching and picking up a target. They controlled the robot by classifying 4 sets of movements with a total of fourteen behaviours of the robot. One visual stimuli is used to select the behaviour set and the remainder are used to encode the behaviours. The user interface of the system consists of five fixed stimuli symbols (five frequency values), a display for a live video feedback and a display for the current posture of the robot. The task completed with an average success rate of 88%, an average response time of 3.48 s and an average ITR of 27.3 bpm.

Similar research has been carried out using SSVEP stimuli to control robot-like behaviour in [11] and [16] in which these authors again used fixed size and position stimuli symbols with differing frequencies that indicate different directions for the robot to move toward. In [11] the authors controlled a mobile robot by using 3 different SSVEP frequencies by moving forward or turning to the left or right

in order to avoid the obstacles. The stimuli in [16] consisted of four fixed flickering boxes where each frequency was used to command a mobile robot (forward, backward, turn counter-clockwise/clockwise) to navigate the robot through a maze path.

There are two notable studies that have integrated object detection and recognition [3], [20]. In [3], the authors used seven different frequencies to navigate a mobile robot to a storage rack to grasp an object and delivered into a dustbin with an average mean accuracy of 89.4%. The approach employed an AdaBoost algorithm with Haar features to recognise three objects on the rack for subjects to choose. However, the recognised objects were not flickered as stimuli - instead, there were separate fixed stimuli designed with three different frequencies corresponding to each object.

The authors in [20] used SSVEP with a hybrid-mask feature in which a 3D textured model was rendered and flickered on certain scene objects. In this case, three similar cans which are recognised offline. Subjects for this study teleoperated a humanoid robot HRP-2 (located in Japan from Italy) to control the robot to firstly, grasp a can from a table, navigated the robot to a second table where the robot is required to drop the can on a marked target.

In this present work, by taking advantage of the on-board camera on our humanoid robot and the high-performance scene object detection model of [21], we instead use variable BCI stimuli, embedded within the scene video feed. This is achieved by flickering the flexible size detected object pixel regions with differing SSVEP frequencies. This occurs in the real-time as the humanoid robot navigates a natural indoor environment. In contrast to earlier work [3], [10], [11], [16], [20], our stimuli vary both in terms of pixel pattern, size and on-screen position in-conjunction with the changing nature of the environment the robot is navigating through.

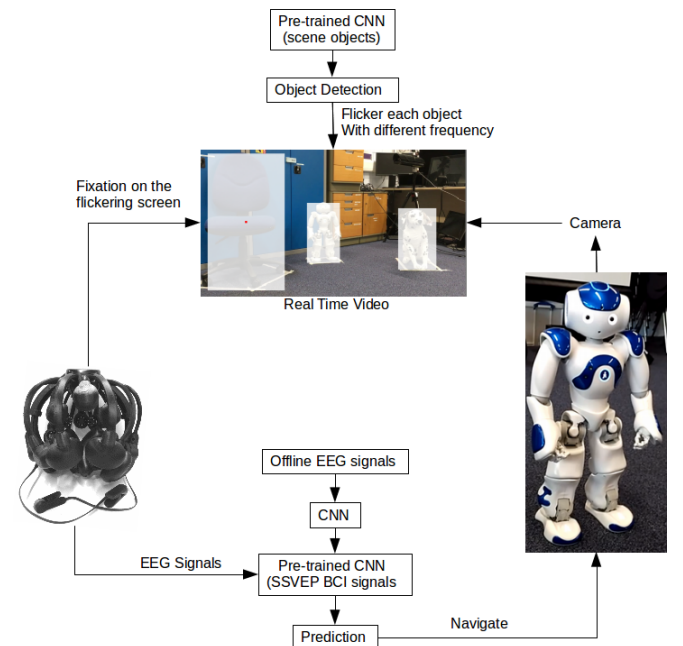


Fig. 1: Overview of the experimental approach proposed.

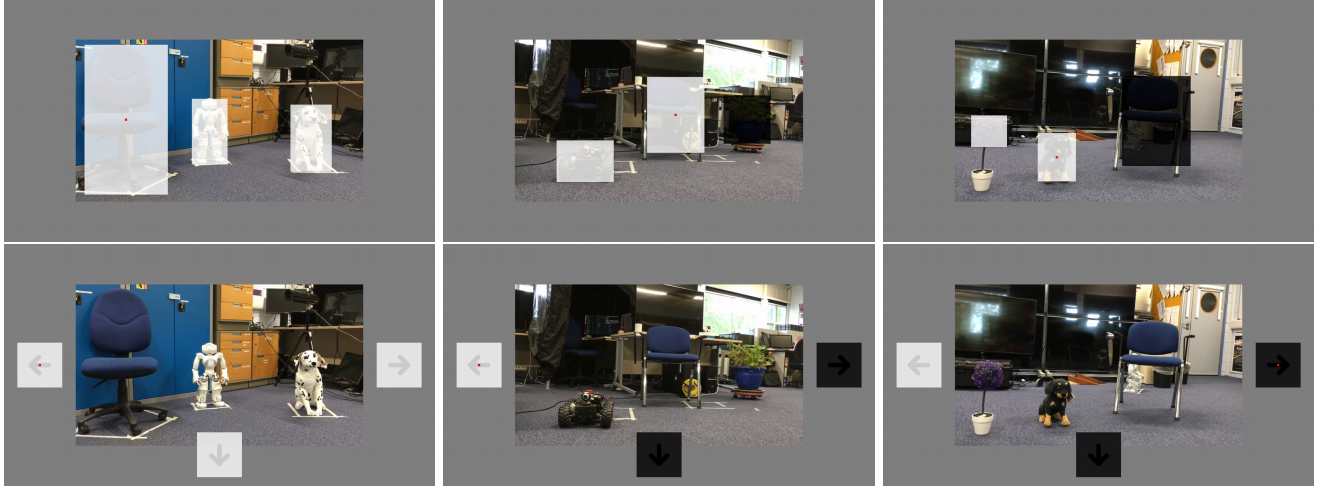


Fig. 2: Variable SSVEP stimuli based on object detection flickering (top row) to navigate the robot towards the object and navigational arrow flickering (bottom row) to move the robot facing a new environmental scene.

III. METHODOLOGY

In this section, we present the four primary experimental components; variable BCI stimuli, streaming dry-EEG signals, EEG signal classification and robot navigation. The overall setup and data flow of the experiment is shown in Figure 1.

We use the on-board camera to stream video from the natural environment to a monitor display in front of the BCI subject (user). Using the CNN-based object detection model of [21], detected scene objects are identified and flickered with a unique on-screen SSVEP frequency (from set: 10Hz, 12Hz, 15Hz). EEG signals from the subject are streamed using the dry-EEG headset whilst they fixate on a flickering on-screen scene object. A CNN pre-trained on an *a priori* offline dataset is then used to infer the class of the EEG signals in real-time. This prediction is used to navigate the robot towards the corresponding scene object the subject is fixated upon.

A. Variable BCI Stimuli

In order to translate the cortical signals, we use SSVEP as the neurophysiological brain response for subjects. The stimuli are embedded into the real-time video streaming from the on-board robot camera (RGB colour, resolution: 1280×960). Based on pre-trained object detection, we flicker the on-screen display frequency of objects by rendering black/white polygon boxes on top of the objects with display frequency modulations of 10, 12 and 15 Hz [15].

In this work, we employ the pre-trained Single Shot MultiBox (SSD) Object Detector CNN [21]. This CNN was trained by using the 12 objects class from the COCO dataset [22]. We present the stimuli using [23] on a 60Hz refresh rate LCD monitor.

The teleoperation interface display alternates between this detected object flickering and navigational arrow flickering one after another as illustrated in Figure 2. The additional use of the navigational arrow stimuli enables the subject to navigate the robot when there is no new object detected

within the scene, for example, when the robot is too close to the previously subject (user) selected object.

B. Dry-EEG Signal Streaming

We use the Cognionics Quick-20 20-channel dry-EEG headset to stream the cortical signals from three healthy subjects (S01, S02, S03) whilst each subject is sat in front of the variable SSVEP stimuli. The dry-EEG headset provides 19 channels and A2, reference and ground as in Figure 3 with a 10-20 compliant sensor layout (international standard for reproducible sensor placement across different EEG experiments [7]).

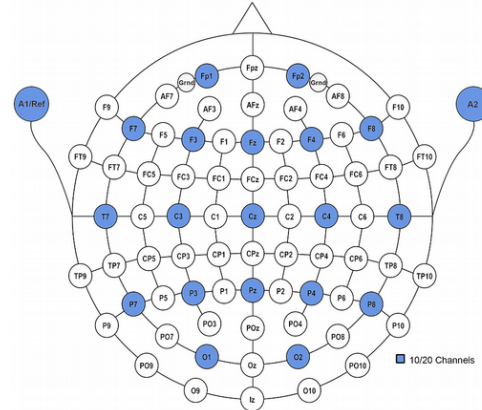


Fig. 3: The location of the electrodes of the dry-EEG headset highlighted in blue within the 10/20 EEG reference model.

This portable and wireless headset is straightforward and easy-to-use as it does not require neither any skin preparation prior to use or conductive gel (as wet-EEG).

During the experiments, we stream the signals over nine sensors; parietal and occipital cortex (P7, P3, Pz, P4, P8, O1 and O2) [4], [6], [15], frontal centre (Fz) and A2 reference at 500 Hz sampling rate for three seconds per trial. The odd numbers represent the left hemisphere of the brain, whilst the even represent the right hemisphere (Figure 3).

The dry-EEG headset requires proprietary data acquisition software, used to measure impedance values before use to ensure optimal-quality dry-EEG signals. In addition, it streams the data from the headset to a computer and allows networked data access to send the data streaming over the network (between two different computers, for example).

C. EEG Signals Classification

To decode the dry-EEG signals efficiently in order to ensure effective teleoperation of the robot, we use our deep CNN architecture of [17]) (see reference for more details) for signal to object/motion label classification.

During the offline experiments, subjects fixate to one of the flickering stimuli. The cortical brain signals from each subject are collected for 40 experimental trials per SSVEP class to form the offline *a priori* training sets or training the CNN model per subject (offline calibration).

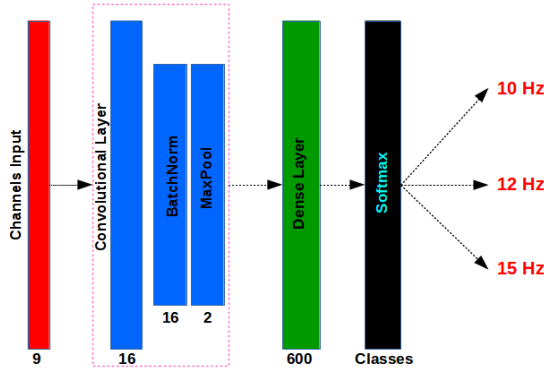


Fig. 4: The 1D CNN architecture used to classify the EEG signals for both offline dataset and real-time experiments (SCU, highlighted in pink).

We train a SSVEP Convolutional Unit (SCU) CNN architecture [17], comprising of a 1D convolutional layer, batch normalization and max pooling (as detailed in Figure 4) by using the offline prior experimental datasets. We first bandpass filter the incoming signals between 9 to 100 Hz in order to reduce undesired high or low frequencies that are not of interest in this work. The filtered signals which consist of nine input channels are transformed by using a large initial convolutional filter to capture the frequencies we are interested in classifying the dry-EEG data. The SCU CNN model is trained using backpropagation with stochastic gradient descent [24].

For this training, the key hyperparameters, initially chosen via a grid-search over a validation set, are L2 weight decay scaling 0.004, dropout level 0.5, convolution kernel size 1×10 , kernel stride 4, maxpool kernel size 2, categorical cross entropy as the optimisation function, ADAM gradient descent algorithm [25] and ReLU as the activation function on all hidden layers.

D. Robot Navigation

The experiment begins with the robot facing a scene containing objects which are detected to generate on-screen SSVEP stimuli pixel regions as previously outlined. The

subject (teleoperation user) fixates on one particular object from which robot navigation is performed using the high level mobility functions of the NAO humanoid robot platform (Figure 2), based on the decoding of the corresponding SSVEP signals by the pre-trained SCU CNN model (Section III-C).

Once these BCI signals are classified as a selected scene object by the subject (user), we then calculate the required robot motion trajectory. As we cannot acquire depth information directly from the monocular camera on the robot, we acquire the distance and the angle of view of the chosen object following the photogrammetric approach of [26]. As such, the distance of the object Z can be calculated as:

$$Z = f' \frac{Y}{y}, \quad (1)$$

where Z is the distance in metres, f' is the focal length (pixels), Y is the object height (metres) and y is the object height in the image (pixels) f' as:

$$f' = H \frac{f}{h}, \quad (2)$$

where H is height of the image (pixels), f is the focal length in metres and h is the sensor height in metres.

The angle of view (AoV) of the object from the camera in radian based on the horizontal position x of the image in pixel can be calculated as follows:

$$AoV = \frac{x}{f'}. \quad (3)$$

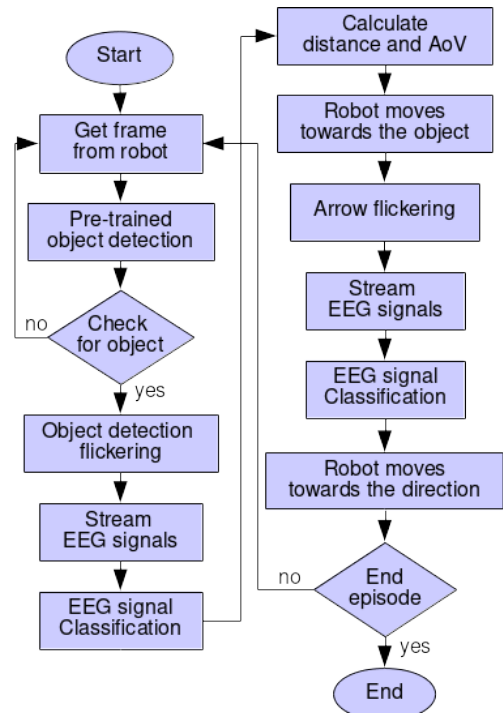


Fig. 5: Flowchart of real-time robot navigation.

When the robot navigates within a given distance and angle trajectory of the subject-selected object, the BCI on-screen interface display alternates to the navigational arrow display (left, right, backwards – Figure 2) using the specific SSVEP frequencies of 10, 12 and 15 Hz. These frequencies intend to facilitate robot motion at 90 degree turns left/right or a 180 degree about turn. Subjects similarly attend to one of these SSVEP stimuli which, once decoded by the SCU CNN model, facilitate general robot motion in the environment until further scene objects are detected within the scene traversal. A flow diagram operation of the real-time experimental teleoperation of the NAO robot through the environment in this alternating object-stimuli and navigational-stimuli manner is presented in Figure 5.

The experimental navigation plan used during the real-time experiments presented in this study is shown in Figure 6. Under these conditions, we repeat the experimental episode five times per subject to demonstrate the repeatability and robustness of our approach.

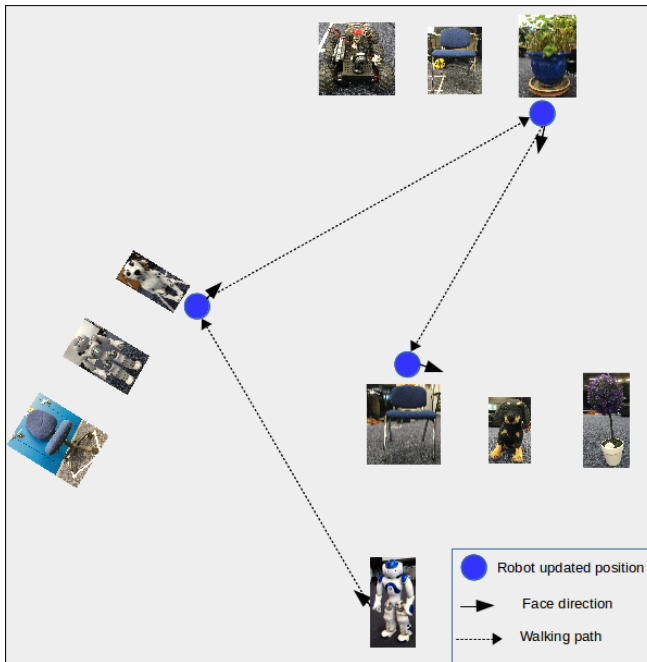


Fig. 6: Navigation plan for real-time experimentation.

IV. RESULTS AND DISCUSSION

In this section, we present the results from the offline classification and the real-time experiment classification using the metrics of classification accuracy and Information Transfer Rate (ITR) in bits per minute (bpm).

A. Offline Statistical Performance

The result for the classification accuracy and the ITR of the offline experiment are presented in Table I. ITR is the speed of BCI in term of bit rate transfer which is the amount of the information throughput by a system per minute [2].

ITR is a suitable BCI performance metric, as a high ITR is dependent upon high accuracy. The ITR is calculated as in [20]:

Subject	S01	S02	S03
Accuracy	0.96±0.02	0.80±0.09	0.75±0.12
ITR (bpm)	23.90±0.72	12.16±0.95	9.61±0.89

TABLE I: Mean accuracy and ITR with standard deviation for offline classification over 10-fold cross validation.

$$ITR = \frac{B}{T}, \quad (4)$$

where T is the time taken to classify a trial in minutes and B is the bits per trial:

$$B = \log_2(N) + P \log_2(P) + (1 - P) \log_2\left(\frac{1 - P}{N - 1}\right), \quad (5)$$

where N is the number of possible selections (N = 3) and P is the correct selection accuracy.

For the offline experiment, the time taken is based on the total flickering time per trial (3 seconds) plus the average of time the classifier takes to train and classify a trial. The data collected during the offline experimental phase is used to train the model for real-time experimentation. However, in order to demonstrate statistical performance of our SCU CNN architecture on this task, we present mean accuracy over 10-fold cross validation per subject. This is used as the P value to calculate the value for B (Equation 5).

B. On-line Real-time Performance

The results of the on-line experimental phase are presented in Table II where we can see the correlation between the results from both experiments. Overall, the results demonstrate extremely high accuracy for all of the subjects tested.

Subject	S01	S02	S03
Experiment 1	1.00	1.00	1.00
Experiment 2	1.00	0.83	0.50
Experiment 3	0.83	0.83	0.83
Experiment 4	0.83	1.00	1.00
Experiment 5	0.83	0.67	0.67
Mean Accuracy	0.90±0.08	0.87±0.12	0.80±0.19
Mean ITR (bpm)	16.8±0.10	15.6±0.12	13.2±0.16

TABLE II: Accuracy for each experiment and mean accuracy and ITR with standard deviation for real-time classification.

Our results demonstrate a strong statistical performance, with a mean accuracy of 0.85 across all subjects. This is comparable to [10] which obtained 0.88 accuracy, despite our work using a variable SSVEP stimuli. As ITR represents the speed of the real-time information transfer from stimuli to motion command generation, the time taken is measured from the beginning of a stimuli flashing until getting a prediction. We can thus improve the ITR further via reducing the flickering time during the real-time experiment.

Figure 7 represents per-class confusion matrices for the real-time classification and highlights overall good accuracy across all classes for all three subjects (users), although the middle class (12 Hz) is more difficult to classify than the rest of the classes.

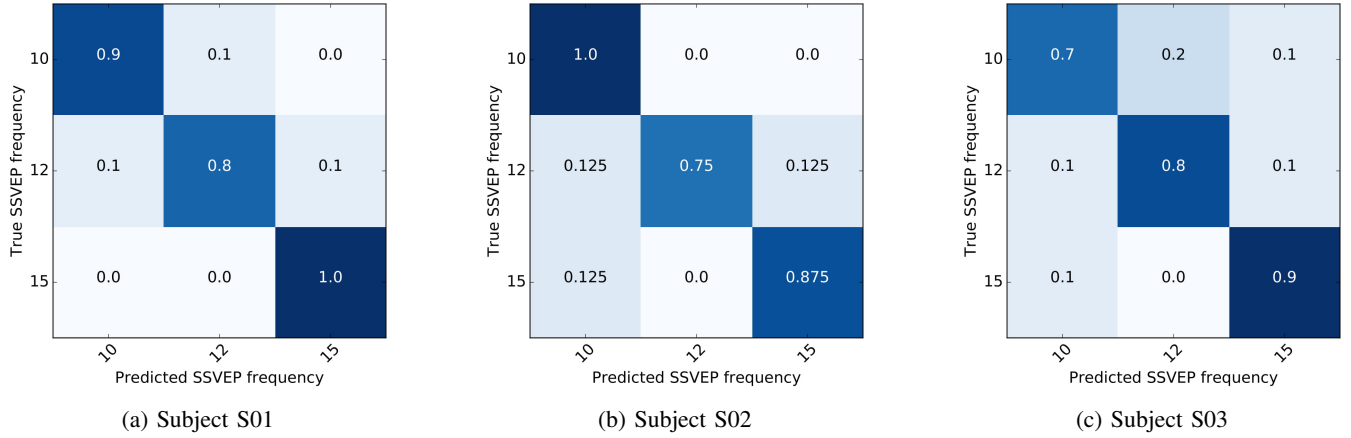


Fig. 7: Confusion matrices for the classification of real-time EEG signals during the robot navigation (maximal result being *accuracy* = 1.0 in the matrix diagonals).

Figure 8 illustrates the real-time experimental environment such as the view from the robot and the robot approaching an object. The angle of direction from the robot to the selected objects can vary from one experiment to another, because the calculation of distance and direction is based on the bounding box from the object detection and the angle of view of an object on the plane. The detected bounding box for the scene object can vary and the angle of view of an object can change with the slightest movement of either the robot or the robot head (where the camera is located).

V. CONCLUSION

In this work, we present a number of novel contributions spanning the use of variable SSVEP stimuli (pattern, size, shape) as an enabler to future telepresence BCI applications in a real-world natural environment. We integrate recent advances in the use of deep CNN architectures for both scene object detection and dry-EEG bio-signal decoding. Within this context, we develop a novel SSVEP interface to flicker the on-screen frequency of naturally occurring objects detected within the scene, as seen from the on-board camera of a teleoperated robot, and decode these dry-EEG brain-based bio-signals based on the frequency of the visual fixation detected to navigate the robot within the scene. Uniquely, we train and utilize a common CNN model (SCU, Figure 4) for use with SSVEP stimuli that vary in size, on-screen position and internal (pixel pattern) throughout the duration of the experiment, significantly advancing such decoding generality against prior work in the field [3], [20]. Our evaluation is presented in terms of accuracy and ITR, both on the *a priori* experimental training set used for the off-line training phase (via cross validation) and the on-line real-time teleoperated navigation of a humanoid robot through a natural indoor environment. The introduction of these highly novel and variable BCI SSVEP stimuli, based on scene object occurrence, demonstrates adaptable BCI-driven robot teleoperation within a natural environment (without scene markers and alike). Strong statistical classification performance is observed, comparable to and often exceeding

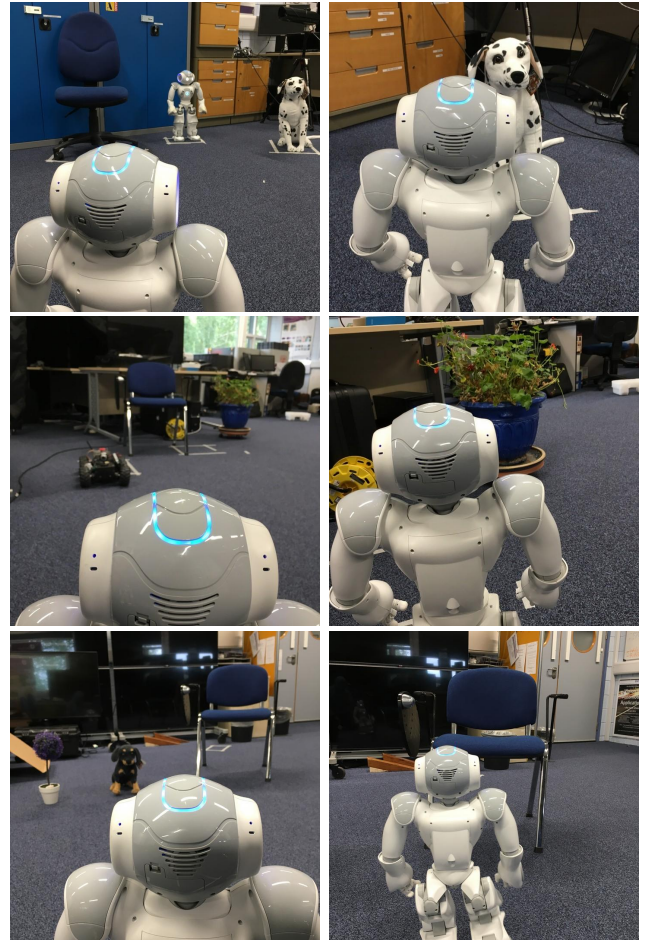


Fig. 8: Sample of humanoid robot navigation during a real-time experiment.

those reported in the general BCI literature [10], despite the introduction of the serious challenges associated with variable SSVEP stimuli.

Future work will look to improve generalisation performance over additional test subjects, increase both scene complexity and teleoperative duration as well as considering aspects of robot interaction within the environment.

REFERENCES

- [1] L. Bi, X.-A. Fan, and Y. Liu, "EEG-based Brain-Controlled Mobile Robots: A Survey," *IEEE transactions on Human-Machine Systems*, vol. 43, no. 2, pp. 161–176, 2013.
- [2] R. P. Rao, *Brain-Computer Interfacing: An Introduction*. New York, NY, USA: Cambridge University Press, 2013.
- [3] S. Sheng, P. Song, L. Xie, Z. Luo, W. Chang, S. Jiang, H. Yu, C. Zhu, J. T. C. Tan, and F. Duan, "Design of an SSVEP-based BCI System With Visual Servo Module for a Service Robot to Execute Multiple Tasks," in *IEEE International Conference on Robotics and Automation*. IEEE, 2017, pp. 2267–2272.
- [4] Q. Gao, L. Dou, A. N. Belkacem, and C. Chen, "Noninvasive Electroencephalogram Based Control of a Robotic Arm for Writing Task using Hybrid BCI System," *BioMed Research International*, vol. 2017, 2017.
- [5] J. Minguiñon, M. A. Lopez-Gordo, and F. Pelayo, "Trends in EEG-BCI for Daily-life: Requirements for Artifact Removal," *Biomedical Signal Processing and Control*, vol. 31, pp. 407–418, 2017.
- [6] Y.-P. Lin, Y. Wang, C.-S. Wei, and T.-P. Jung, "Assessing the Quality of Steady-State Visual-Evoked Potentials for Moving Humans using a Mobile Electroencephalogram Headset," *Frontiers in Human Neuroscience*, vol. 8, no. March, pp. 1–10, 2014.
- [7] T. R. Mullen, C. A. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time Neuroimaging and Cognitive Monitoring using Wearable Dry EEG," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2553–2567, 2015.
- [8] M. A. Lopez-Gordo, D. Sanchez-Morillo, and F. Pelayo Valle, "Dry EEG Electrodes," *Sensors*, vol. 14, no. 7, pp. 12 847–12 870, 2014.
- [9] G. Lisi, M. Hamaya, T. Noda, and J. Morimoto, "Dry-wireless EEG and Asynchronous Adaptive Feature Extraction Towards a Plug-and-play Co-adaptive Brain Robot Interface," in *IEEE International Conference on Robotics and Automation*. IEEE, 2016, pp. 959–966.
- [10] J. Zhao, W. Li, X. Mao, H. Hu, L. Niu, and G. Chen, "Behavior-based SSVEP Hierarchical Architecture for Telepresence Control of Humanoid Robot to Achieve Full-body Movement," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 2, pp. 197–209, 2017.
- [11] S. Liu, F. Wang, S. Wu, Y. Zhang, Y. Wei, W. Wu, H. Zhao, and Y. Zhang, "Research of Mobile Robot Control System Based on SSVEP Brain Computer Interaction," in *2018 Chinese Control And Decision Conference*. IEEE, 2018.
- [12] C.-Y. Chiu, A. K. Singh, Y.-K. Wang, J.-T. King, and C.-T. Lin, "A Wireless Steady State Visually Evoked Potential-based BCI Eating Assistive System," in *International Joint Conference on Neural Networks*. IEEE, 2017, pp. 3003–3007.
- [13] S. Naz and N. Bawane, "Recent Trends in BCI based Speller System: A Survey Report," *International Journal of Engineering Science*, vol. 6, no. 7, 2016.
- [14] S. K. Andersen and M. M. Müller, "Driving Steady-State Visual Evoked Potentials at Arbitrary Frequencies using Temporal Interpolation of Stimulus Presentation," *BMC Neuroscience*, vol. 16, no. 1, p. 95, 2015.
- [15] X. Mao, M. Li, W. Li, L. Niu, B. Xian, M. Zeng, and G. Chen, "Progress in EEG-based Brain Robot Interaction Systems," *Computational Intelligence and Neuroscience*, vol. 2017, 2017.
- [16] S.-C. Chen, Y.-J. Chen, I. A. Zaeni, and C.-M. Wu, "A Single-Channel SSVEP-Based BCI with a Fuzzy Feature Threshold Algorithm in a Maze Game," *International Journal of Fuzzy Systems*, vol. 19, no. 2, pp. 553–565, 2017.
- [17] N. K. N. Aznan, S. Bonner, J. D. Connolly, N. A. Moubayed, and T. P. Breckon, "On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks," in *IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2018, pp. 3716–3721.
- [18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [19] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with Convolutional Neural Networks for EEG Decoding and Visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [20] E. Tidoni, P. Gergondet, G. Fusco, A. Kheddar, and S. M. Aglioti, "The Role of Audio-Visual Feedback in a Thought-based Control of a Humanoid Robot: A BCI Study in Healthy and Spinal Cord Injured People," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 6, pp. 772–781, 2017.
- [21] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot Multibox Detector," in *European Conference on Computer Vision*. Springer, 2016, pp. 21–37.
- [22] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," in *European Conference on Computer Vision*. Springer, 2014, pp. 740–755.
- [23] J. W. Peirce, "PsychoPy-Psychophysics Software in Python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.
- [24] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [25] D. P. Kingma and J. Ba, "Adam: A method for Stochastic Optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [26] M. E. Kundegorski and T. P. Breckon, "A Photogrammetric Approach for Real-time 3D Localization and Tracking of Pedestrians in Monocular Infrared Imagery," in *Optics and Photonics for Counterterrorism, Crime Fighting, and Defence X; and Optical Materials and Biomaterials in Security and Defence Systems Technology XI*, vol. 9253. International Society for Optics and Photonics, 2014, p. 92530I.